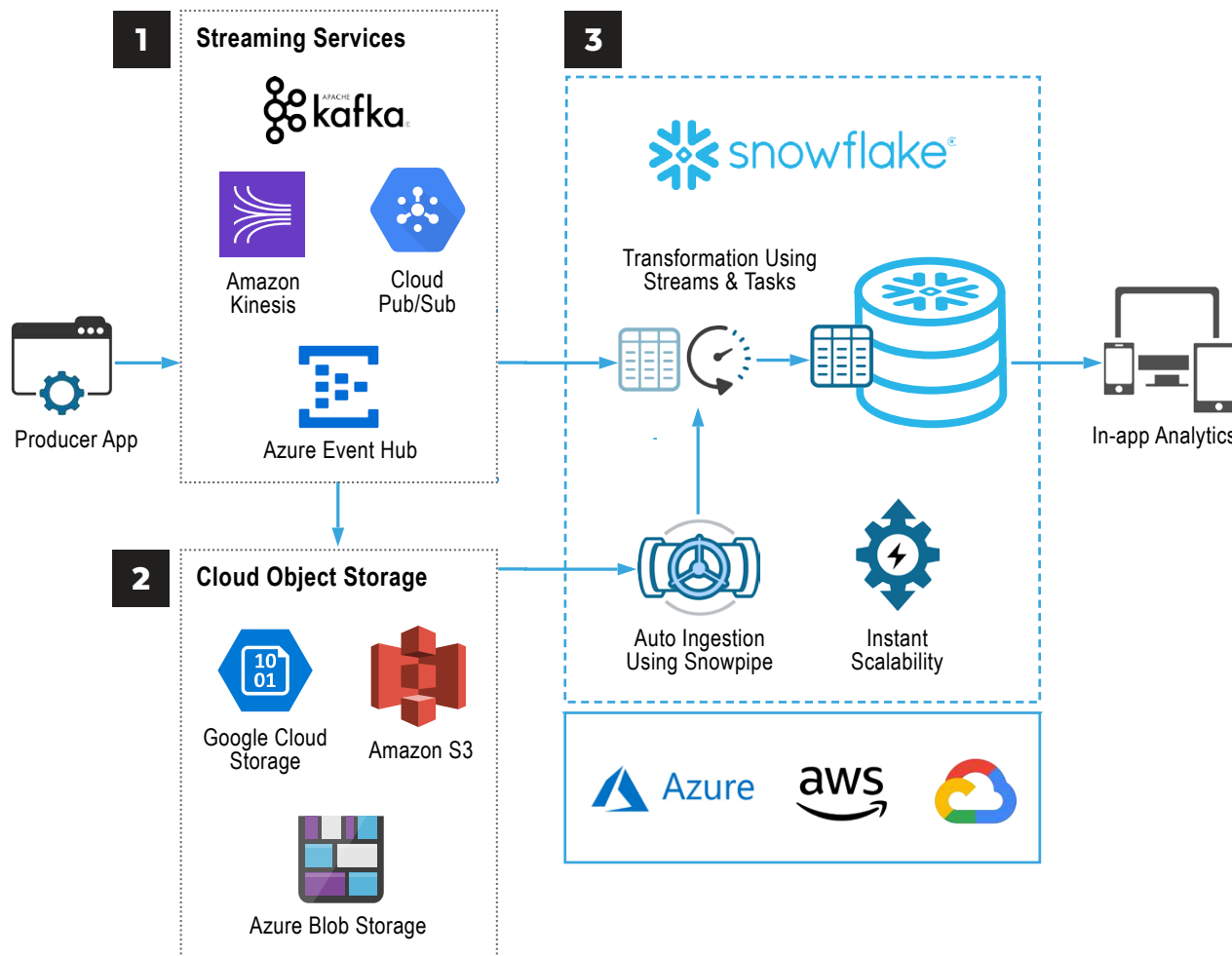


STREAMING DATA STACK REFERENCE ARCHITECTURE

STREAMING DATA STACK



OBJECTIVE

Build data intensive applications that rely on streaming data ingestion and analysis.

DESCRIPTION

- 1** The producer application generates continuous data that the streaming service ingests and buffers to account for data rate differences between the producer and consumers. Depending on the application's needs, Snowflake ingests data directly from the streaming service or via cloud object storage (2).
- 2** In cases where the application requires raw data to persist in cloud object storage, the streaming service processes the raw data and batches it into larger chunks, thus lowering the API storage expenses. When Amazon Kinesis is used as the streaming service, data is staged in cloud object storage before ingestion.
- 3** Snowflake ingests data from the streaming service into a staging table and stores the streamed data for analysis. Its Streams and Tasks features detect data changes and schedule tasks to perform any required transformations. Multiple streams and tasks can be chained to implement a complex data pipeline. Snowpipe with Auto-Ingest automates the data ingestion from cloud object storage.